The EMBO Journal vol.8 no.7 pp.2087 - 2094, 1989

Structural polymorphism of homopurine - homopyrimidine sequences: the secondary DNA structure adopted by a d(GA.CT)$_{22}$ sequence in the presence of zinc ions

J.Bernués, R.Beltrán, J.M.Casasnovas and F.Azorín

Grupo de Química Macromolecular, Centro de Investigaciòn y Desarrollo CSIC, ETSEIB Diagonal 647, Barcelona 08028, Spain

Communicated by J.Subirana

In this paper, we have analysed the conformational behaviour shown by the homopurine - homopyrimidine alternating d(GA.CT)$_{22}$ sequence cloned into SV40. Our results show that, in the presence of zinc ions, the d(GA.CT)$_{22}$ sequence adopts an altered secondary DNA structure (*H-DNA) which differs from either B-DNA or H-DNA. Formation of *H-DNA is facilitated by negative supercoiling and does not appear to require base protonation, since it is induced at neutral pH by ~0.4 mM $ZnCl_2$. The patterns of OsO_4 and DEPC modification obtained in the presence of zinc are compatible with a homopurine - homopurine - homopyridimine triplex, though other structural models for *H-DNA are also possible. The hypersensitivity to S1-cleavage of the d(GA.CT)$_{22}$ sequence is reinterpreted in terms of the equilibria between the B-, H- and *H-forms of the sequence. These results reveal the high degree of structural polymorphism shown by homopurine - homopyrimidine sequences. Its biological relevance is discussed.

Key words: *H-DNA/Pur - Pyr sequences/S1-hypersensitivity/triplex/zinc

Introduction

Depending on the precise nucleotide sequence and environmental conditions, DNA in solution can exist under a variety of different structural conformations. Recently, much attention has been devoted to the characterization of the conformational behaviour depicted by homopurine - homopyrimidine stretches in DNA. Homopurine - homopyrimidine sequences are frequently located at the 5′ flanking region of many eukaryotic genes (Larsen and Weintraub, 1982; Nickol and Felsenfeld, 1983; Schon *et al.*, 1983; Elgin, 1984; Evans *et al.*, 1984; McKeon *et al.*, 1984; Kilpatrick *et al.*, 1986) and they are hypersensitive to S1 nuclease and other single-stranded specific nucleases, in naked DNA as well as in chromatin (Larsen and Weintraub, 1982; Weintraub, 1983). Homopurine - homopyrimidine sequences are also frequent in sites involved in genetic recombination (Davis *et al.*, 1980; Hentschel, 1982; Moos and Gallwitz, 1983; Htun *et al.*, 1984; Wohlrab *et al.*, 1987). In general, preferential cleavage by single-stranded nucleases has been interpreted as indicative of the adoption of a non-B secondary DNA structure and several models have been proposed to account for the S1-hypersensitivity of these sequences (Cantor and Efstratiadis, 1984; Pulleyblank *et al.*, 1985; Kohwi-Shigematsu and Kohwi, 1985; Lyamichev *et al.*, 1985, 1986; Evans and Efstratiadis, 1986).

The homopurine - homopyrimidine alternating d(GA.CT)$_n$ sequence constitutes a significant proportion of the simple repeating DNA sequences found in eukaryotic genomic DNA (Tautz and Renz, 1984; Gross and Garrard, 1986). d(GA.CT)$_n$ sequences are known to undergo transition to an unwound, S1-hypersensitive DNA conformation in response to protonation and decreasing supercoiling (Pulleyblank *et al.*, 1985; Lyamichev *et al.*, 1985). Other homopurine - homopyrimidine sequences, including d(G.C)$_n$, have been found to undergo similar structural transitions (Lyamichev *et al.*, 1987). On the basis of the unwinding obtained from the transition, a model for the altered conformation adopted by d(GA.CT)$_n$ sequences in supercoiled DNA was proposed (Lyamichev *et al.*, 1986). In this model, called H-form DNA, about half of the homopyrimidine strand folds back upon itself forming Hoogsteen pairs with the purines, which results in a triplex containing CGC^+ and TAT base triads. On the other hand, the second half of the homopurine strand forms a single-stranded loop, which accounts for the S1-hypersensitivity of the sequence. This model is supported by abundant data regarding the chemical reactivity of the altered DNA conformation (Vojtišková and Palecek, 1987; Hanvey *et al.*, 1988; Johnston, 1988; Kowhi and Kowhi-Shigematsu, 1988; Vojtišková *et al.*, 1988; Voloshin *et al.*, 1988). However, some features of the S1-cleavage pattern, and in particular the S1 sensitivity of the polypyrimidine strand, are not easily interpreted in terms of the H-form model. Other models have also been proposed, which mainly address the question of the S1 cleavage pattern of this sequence. Pulleyblank *et al.* (1985) have suggested that the homopurine - homopyrimidine alternating d(GA.CT)$_n$ sequence might adopt a right-handed double-stranded DNA conformation in which usual dA:dT Watson - Crick base pairs alternate with Hoogsteen ^{syn}dG:dC^+ base pairs. The S1 sensitivity of the sequence will arise from the peculiar configuration of the phosphodiester backbone, which will be recognized and cleaved by the nuclease. Protonation of the cytosine residues is, however, a common principal feature of both models. On the other hand, Evans and Efstratiadis (1986) have proposed that in homopurine - homopyrimidine sequences the two strands have different backbone conformations (heteronomous DNA).

In this paper, the conformational behaviour shown by a d(GA.CT)$_{22}$ sequence cloned into SV40 has been analysed. In contrast to what would be expected for a B-to-H transition, our results show that the d(GA.CT)$_{22}$ insert is as hypersensitive to S1-cleavage at neutral pH as it is at acidic pH. Both strands are cleaved by S1 and the cleavage pattern does not change much with pH, particularly that of the homopyridimine strand. On the other hand, the patterns of DEPC

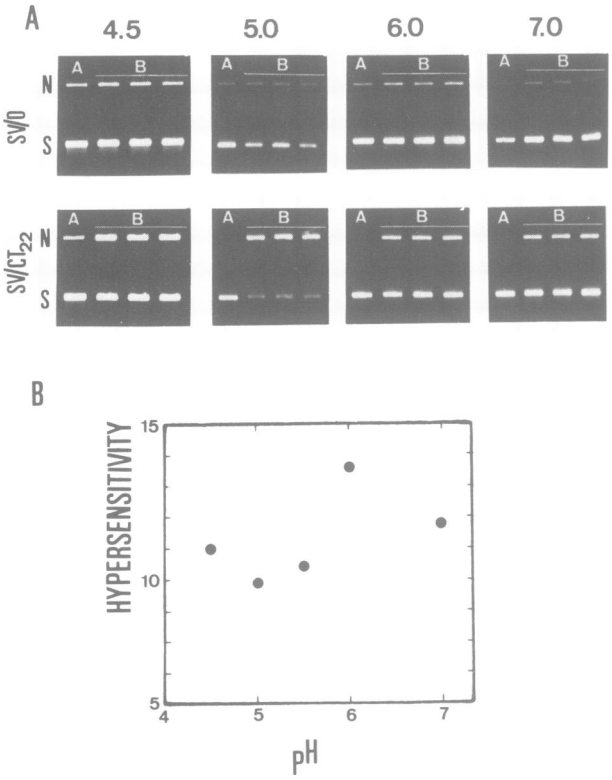

Fig. 1. Hypersensitivity to S1 cleavage of the d(GA.CT)$_{22}$ sequence as a function of pH. (**A**) 1% agarose−TPE gel electrophoretic analysis of purified SV/O and SV/CT$_{22}$ DNAs treated with S1 nuclease (**lanes B**) at the pH values indicated. Digestion with S1 was carried out at 15°C for 30 min at enzyme/DNA ratios of: 0.1 U/μg (pH 4.5); 10 U/μg (pH 5.0 and 6.0) and 50 U/μg (pH 7.0). **Lanes A** correspond to samples untreated with S1. All experiments were performed in triplicate. N, nicked; S, supercoiled. (**B**) Quantitation of the results shown in (A). Hypersensitivity to cleavage by S1 is expressed as the ratio of the percentage of nicked SV/CT$_{22}$ DNA molecules divided by the percentage of nicked SV/O DNA molecules. Each point corresponds to the average of several experiments carried out independently at enzyme/DNA ratios of: 0.1−1 U/μg (pH 4.5); 1−10 U/μg (pH 5.0, 5.5 and 6.0) and 10−50 U/μg (pH 7.0). All experiments were performed with negatively supercoiled DNA ($-\bar{\sigma} = 0.05$).

or OsO_4 modification obtained at acidic pH agree quite well with the H-form model. However, when modification with either DEPC or OsO_4 is performed in the presence of zinc ions (the counter-ion normally used for S1 cleavage), a different pattern of modification is obtained which is not consistent with the H-form model, indicating that, in the presence of zinc ions, the d(GA.CT)$_{22}$ sequence adopts a non-B, non-H secondary DNA conformation, that we call *H-DNA. *H-DNA occurs at neutral pH and is stabilized by negative supercoiling. These results indicate that, depending upon the exact environmental conditions−presence or absence of zinc ions, proton concentration, etc.−the alternating d(GA.CT)$_n$ sequence can adopt different structural conformations. The high degree of conformational plasticity shown by d(GA.CT)$_n$ sequences, and in general homopurine−homopyrimidine sequences, might be on the basis of the discrepancies between the different models proposed for its altered DNA conformation(s).

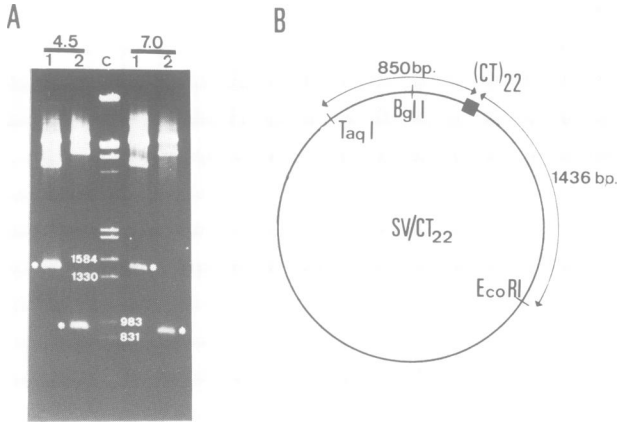

Fig. 2. Mapping of S1-hypersensitive sites in SV/CT$_{22}$ DNA. (**A**) 1% agarose−TBE gel electrophoretic analysis of SV/CT$_{22}$ DNA treated with S1 nuclease at pH 4.5 (25 U/μg) and pH 7.0 (100 U/μg) as indicated in Materials and methods and then subjected to either *Eco*RI (**lanes 1**) or *Taq*I (**lanes 2**) restriction. **Lane C** corresponds to λ-DNA *Eco*RI + *Hin*dIII restriction fragments used as mol. wt markers. The size in base pairs of selected markers is indicated. (**B**) Size of the *Eco*RI and *Taq*I fragments (indicated by an asterisk in A) expected if the main S1-hypersensitive sites of SV/CT$_{22}$ DNA were confined to the d(GA.CT)$_{22}$ sequence.

Results

The d(GA.CT)$_{22}$ sequence is preferentially cleaved by S1 nuclease at neutral pH

The construction and genomic organization of SV/CT$_{22}$ and SV/O recombinants have been described in detail elsewhere (Casasnovas *et al.*, 1987). SV/CT$_{22}$ is an SV40 recombinant that carries a d(GA.CT)$_{22}$ insertion at the *Hpa*II site. SV/O differs from SV/CT$_{22}$ in that it does not contain any alternating repeated sequence; otherwise its genomic organization is identical to that of SV/CT$_{22}$.

The sensitivity of the d(GA.CT)$_{22}$ sequence to cleavage by S1 nuclease at different pH values was analysed by determining the relative nicking of purified SV/CT$_{22}$ DNA with respect to purified SV/O DNA (Figure 1). It is known that the enzymatic activity of S1 nuclease is highly dependent on pH, being maximal at around pH 4.5 and decreasing progressively as pH increases (Vogt, 1980). Therefore the percentage of nicking induced by S1 nuclease is dependent on pH. However, conditions can be found by just raising the amount of enzyme concomitantly with the decrease in enzymatic activity, where SV/CT$_{22}$ DNA shows approximately the same hypersensitivity to S1-cleavage regardless of pH (Figure 1A). Actually the relative nicking of SV/CT$_{22}$ versus SV/O remains constant or increases slightly as pH increases (Figure 1B), indicating that SV/CT$_{22}$ DNA is as hypersensitive to cleavage by S1 at neutral pH as it is at acidic pH.

That S1 nuclease cleavage of SV/CT$_{22}$ DNA occurs predominantly within the d(GA.CT)$_{22}$ insert was determined by mapping the main S1-hypersensitive site(s) in SV/CT$_{22}$ DNA (Figure 2). Purified, negatively supercoiled SV/CT$_{22}$ DNA was digested with S1 nuclease at either acidic pH (pH 4.5) or neutral pH (pH 7.0) and then subjected to restriction with either *Eco*RI endonuclease (Figure 2A, lanes 1) or *Taq*I endonuclease (Figure 2A, lanes 2). The

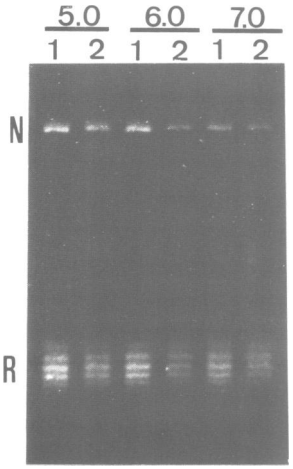

Fig. 3. The S1-hypersensitivity of SV/CT$_{22}$ DNA depends on supercoiling. Relaxed SV/CT$_{22}$ DNA was treated with S1 nuclease at either, pH 5.0 (10 U/μg), pH 6.0 (10 U/μg) or pH 7.0 (50 U/μg) (**lanes 2**) and resolved in a 1% agarose−TBE gel run in the presence of 1.25 μg/ml of the intercalator chloroquine phosphate. **Lanes 1** show the corresponding untreated samples. N, nicked; R, relaxed.

restriction pattern corresponding to the DNA treated with S1 at neutral pH turns out to be identical to that obtained when the DNA was subjected to S1-cleavage at acidic pH (Figure 2A), indicating that at both pHs, S1 is recognizing the same hypersensitive site(s) in SV/CT$_{22}$ DNA. If the main S1-hypersensitive site(s) of SV/CT$_{22}$ DNA were confined to the d(GA.CT)$_{22}$ sequence, digestion with *Eco*RI or *Taq*I endonuclease of the S1-treated SV/CT$_{22}$ DNA should result in the production of DNA fragments of ~1436 bp and ~850 bp in length respectively (Figure 2B), as is observed both at pH 4.5 and pH 7 (Figure 2A).

The S1 hypersensitivity of the d(GA.CT)$_{22}$ sequence at neutral pH depends on supercoiling. In fact, relaxed SV/CT$_{22}$ DNA is not being preferentially nicked by S1 nuclease either at pH 5.0, 6.0 or neutral pH (Figure 3). No significant differences are detected when the percentage of nicked molecules after S1 treatment of relaxed SV/CT$_{22}$ DNA (Figure 3, lanes 2) is compared with the percentage of molecules that were already nicked before S1 treatment (Figure 3, lanes 1). On the other hand, when negatively supercoiled SV/CT$_{22}$ DNA ($-\bar{\sigma} = 0.05$) was treated with S1 nuclease under the same experimental conditions, a very significant S1-induced nicking was observed (Figure 1). Actually, the S1-induced nicking of relaxed SV/CT$_{22}$ DNA at any pH is approximately of the same magnitude as the S1-induced nicking observed with negatively supercoiled SV/O DNA (Figure 1), indicating that nicking of relaxed SV/CT$_{22}$ DNA is unspecific.

These results show that, at neutral as well as at acidic pH, the homopurine−homopyrimidine insert of SV/CT$_{22}$ DNA is being preferentially recognized and cleaved by S1 nuclease, indicating that the d(GA.CT)$_{22}$ sequence also adopts an altered S1-sensitive DNA conformation at neutral pH, which is stabilized by negative supercoiling. In contrast to what would be expected for a B-to-H transition, the hypersensitivity of the sequence is practically identical at neutral and at acidic pH. Moreover, the negative specific linking difference ($-\bar{\sigma}$) of the DNA used in these experiments was

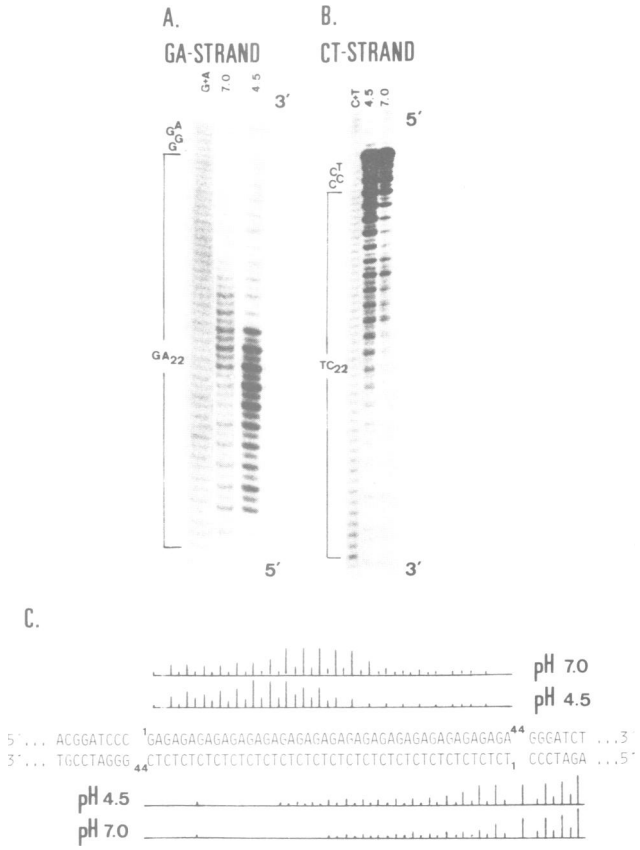

Fig. 4. Patterns of S1 nicking within the d(GA.CT)$_{22}$ sequence. SV/CT$_{22}$ DNA was treated with S1 nuclease at either pH 4.5 (0.08 U/μg) or pH 7.0 (15 U/μg) and processed as described in Materials and methods. (**A**) Nicking pattern of the homopurine strand. (**B**) Nicking pattern of the homopyrimidine strand. (**C**) Summary of the results shown in (A) and (B). The height of each bar is approximately proportional to the sensitivity of the corresponding phosphodiester bond normalized with respect to the most sensitive in the lane. G + A and C + T correspond to sequencing ladders obtained according to Maxam and Gilbert (1980).

of 0.05, which is well below the threshold required to stabilize the H-form at pH 7.0 (Lyamichev *et al.*, 1985).

The pattern of S1 nicking within the d(GA.CT)$_{22}$ sequence does not change much with pH

Figure 4 shows the patterns of S1 nicking within the d(GA.CT)$_{22}$ insert obtained at acidic and neutral pH. From the results presented in Figure 4, the following conclusions can be drawn.

(i) Both strands are cleaved by S1 nuclease. From the type of experiments described in Figure 4, it is difficult to make quantitative conclusions. However, the hypersensitivity of the homopyrimidine strand does not seem to be very different from that of the homopurine strand.

(ii) The patterns of nicking on both strands show a clear dinucleotide repeat. Not all base-steps are equally sensitive to S1-cleavage. In the homopurine strand GpA over ApG steps are being preferentially hit by S1 nuclease. Similarly, in the homopyrimidine strand CpT linkages are more sensitive to S1 attack than TpC linkages.

(iii) S1 nicking shows a certain degree of directionality. On both strands, S1 nicking is more frequent at the 5′ region

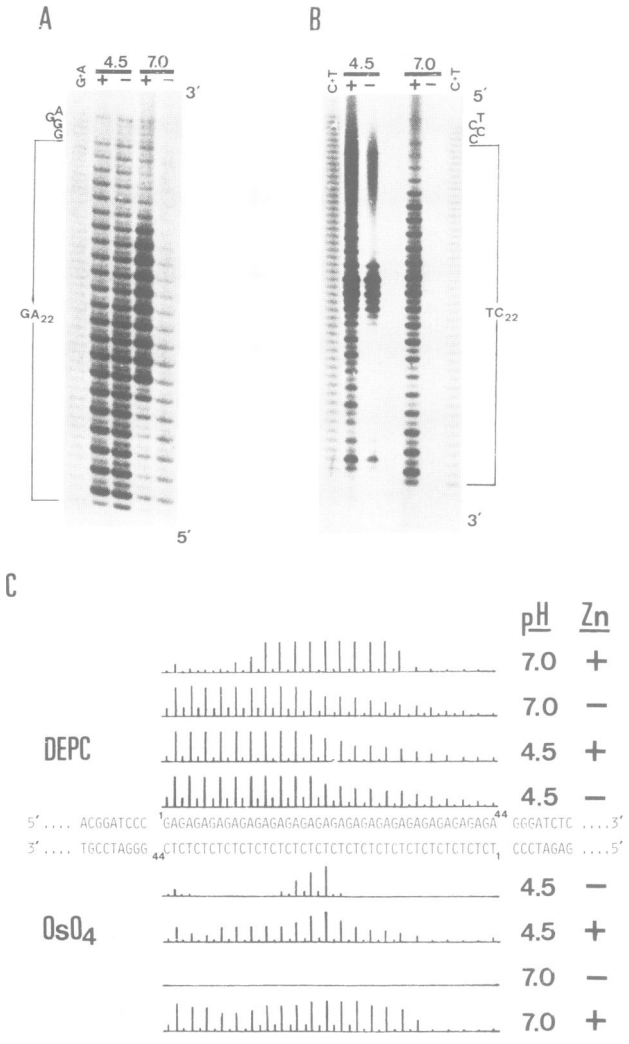

Fig. 5. Patterns of chemical modification within the d(GA.CT)$_{22}$ sequence. SV/CT$_{22}$ was treated with either DEPC or OsO_4 at pH 4.5 and pH 7.0 in the absence (lanes −) or presence (lanes +) of 4 mM $ZnCl_2$, and then processed as indicated in Materials and methods. (**A**) DEPC modification of the homopurine strand. (**B**) OsO_4-pyridine modification of the homopyrimidine strand. (**C**) Summary of the results shown in (A) and (B). The height of each bar is approximately proportional to the extent of modification of the corresponding base normalized with respect to the most reactive in the lane.

of the sequence than at its 3′ region. Particularly, S1 nicking of the homopyrimidine strand is maximal at the 5′ flanking region of the d(GA.CT)$_{22}$ sequence and decreases progressively on the 5′ to 3′ direction. The situation is slightly different on the homopurine strand. In this case, though the 5′ region is always more sensitive to S1 cleavage than the 3′ region, the GpA steps that are nicked more frequently are localized close to the centre of the sequence.

(iv) The nicking pattern does not change much with pH. The nicking patterns obtained at pH 4.5 are very similar to those obtained at pH 7.0. Both strands are cleaved, the dinucleotide repeat feature is conserved, and the same base steps are preferentially cleaved at both pH values. The 5′ to 3′ directionality of nicking is also conserved, although in this case a significant difference is observed on the pattern of cleavage of the homopurine strand. Nicking of the polypurine strand is more displaced towards the centre at pH 7 than at pH 4.5. The highest sensitivity to S1-cleavage at neutral pH is found at GpA steps 17−25, while at acidic pH it is found at GpA steps 13−19. On the other hand, nicking of the homopyrimidine strand is almost identical at both pHs.

The pattern of chemical modifications by DEPC and OsO_4 changes drastically in the presence of zinc ions

Diethylpyrocarbonate (DEPC) and osmium tetroxide (OsO_4) have been widely used as probes for DNA structure. Neither one has been found to modify regular right-handed B-DNA. DEPC, carbethoxylates purines at their N-7 position, mainly adenines (A > G). On the other hand, OsO_4 is specific for pyrimidines. In the presence of pyridine, OsO_4 adds to the C5=C6 double bond of pyrimidines, principally those of thymines (T ⩾ C). Both DEPC and OsO_4 react much more with single-stranded DNA than with double-stranded DNA and they have been used as site-specific probes for left-handed Z-DNA and cruciforms (Lilley and Palecek, 1984; Lukasova *et al.*, 1982; Galazka *et al.*, 1986; Herr, 1985; Johnston and Rich, 1985). They have also been extensively used by several groups to study structural transitions at homopurine−homopyrimidine sequences (Vojtišková and Palecek, 1987; Hanvey *et al.*, 1988; Johnston, 1988; Vojtišková *et al.*, 1988; Voloshin *et al.*, 1988).

The pattern of DEPC modification of the d(GA.CT)$_{22}$ sequence that we have obtained at acidic pH (Figure 5A) is similar to those reported by others for the protonated H-form of the d(GA.CT)$_{22}$ sequence (Johnston, 1988; Voloshin *et al.*, 1988; Hanvey *et al.*, 1988). A characteristic hyperreactivity of adenines located in the 5′ half of the d(GA.CT)$_{22}$ sequence is clearly observed. Approximately the same hyperreactivity is detected from adenine 2 to adenine 20 (Figure 5C). Conversely, adenines in the 3′ half become less and less reactive as their distance from the centre of the sequence increases (Figure 5C). A similar pattern of modification, though less intense, is also observed at pH 7.0 (Figure 5A). However, when DEPC modification is carried out at neutral pH in the presence of zinc ions (4 mM) a drastically different pattern of hyperreactivity is obtained (Figure 5A). Hyperreactive adenines are now localized around the middle of the sequence. Adenines 14−30 are strongly modified by DEPC, while adenines at either end of the insert are only slightly modified (Figure 5C). In addition, guanines are less reactive than when the modification was carried out in the absence of zinc ions (Figure 5A,C). The pattern of hyperreactivity obtained when DEPC modification is carried out in the presence of zinc ions, but at acidic pH is similar to that found in the absence of zinc (Figure 5A), although guanines appear to be less reactive in the presence of zinc.

When OsO_4 is used to probe the homopyrimidine strand, a similar structural transition is observed in the presence of zinc ions (Figure 5B). When OsO_4 modification is carried out at pH 4.5 in the absence of zinc, a pattern characteristic of the protonated form of the d(GA.CT)$_{22}$ sequence is obtained. Thymines at the centre of the sequence−from thymine 23 to 25−become hyperreactive (Figure 5C). A significant hyperreactivity is also found at thymine 43 in the 3′ termini of the d(GA.CT)$_{22}$ sequence. No OsO_4 modification is observed at neutral pH (Figure 5B and C). Similar results have been reported by others (Hanvey *et al.*, 1988;

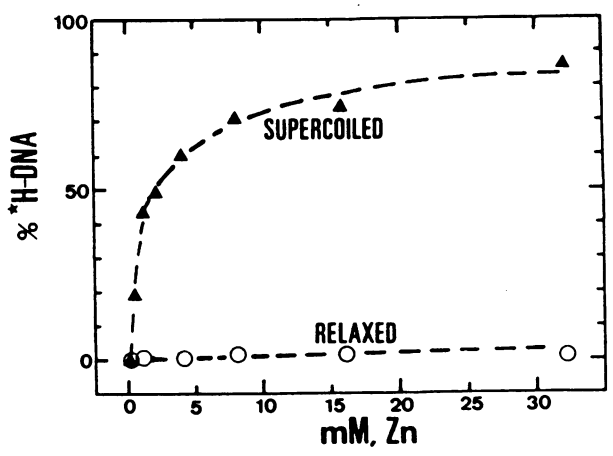

Fig. 6. Stabilization at neutral pH of *H-DNA within the d(GA.CT)$_{22}$ sequence as a function of zinc concentration. Negatively supercoiled ($-\bar{\sigma}$ = 0.05) (▲) or relaxed (○) SV/CT$_{22}$ DNA was modified with OsO_4–pyrimidine at pH 7.0 in the presence of increasing $ZnCl_2$ concentrations. The percentage of *H-DNA is expressed as the percentage of molecules that are cleaved at the d(GA.CT)$_{22}$ sequence by P1 nuclease after OsO_4 modification and linearization (see text for details). Each point corresponds to the average of two independent experiments.

Vojtišková *et al.*, 1988). When zinc ions are present, a completely different pattern of modification is obtained (Figure 5B). Thymines in the 3′ half of the sequence become hyperreactive, though those at the centre are slightly more reactive than the rest, particularly at acidic pH (Figure 5C).

No significant reactivity to DEPC or OsO_4 was observed outside the d(GA.CT)$_{22}$ insert either in the presence or absence of zinc ions.

These results show that in the presence of zinc ions, the d(GA.CT)$_{22}$ sequence adopts a novel structural conformation (*H-DNA) which is different from either the B-form or the protonated H-form of the d(GA.CT)$_{22}$ sequence.

*Induction of *H-form at neutral pH is observed at low concentrations of zinc ions and is facilitated by negative supercoiling*

To determine the zinc dependence of the stabilization of the altered *H-form of the d(GA.CT)$_{22}$ sequence at neutral pH, we took advantage of its observed extensive OsO_4 modification (Figure 5B). It is known that OsO_4-modified sequences are sensitive to cleavage by single-stranded specific nucleases (Lilley and Palecek, 1984; Nejedly *et al.*, 1985; Vojtišková and Palecek, 1987). Therefore, purified negatively supercoiled SV/CT$_{22}$ DNA ($-\bar{\sigma}$ = 0.05) was first modified with OsO_4–pyridine at pH 7.0 in the presence of various concentrations of zinc ions and then digested with P1 nuclease at pH 8.0 after linearization with *Eco*RI. Under these conditions, only those molecules containing the d(GA.CT)$_{22}$ insert in the *H-form will be modified by OsO_4 and thus cleaved by P1. Therefore, the percentage of molecules containing the d(GA.CT)$_{22}$ sequence stabilized in the *H-form at each zinc concentration can be quantitated by simply determining the percentage of molecules that are cleaved by P1 nuclease at the d(GA.CT)$_{22}$ insert after OsO_4 modification and linearization.

In agreement with the results shown in Figure 5B, no P1 cleavage of the OsO_4-treated negatively supercoiled SV/CT$_{22}$ DNA is observed in the absence of zinc ions (Figure 6). Increasing zinc concentration results in a progressive stabilization of the *H-form as judged by P1 cleavage, so that at 32 mM $ZnCl_2$, close to 90% cleavage is obtained (Figure 6). No cleavage was observed when the DNA was incubated at this zinc concentration in the absence of OsO_4. The percentage of molecules cleaved by P1 nuclease increases rapidly between 0 mM and 5 mM $ZnCl_2$ and then it plateaus. At relatively low zinc concentration a significant proportion of the molecules contain the d(GA.CT)$_{22}$ insert stabilized as *H-DNA. At 0.4 mM $ZnCl_2$ ~20% of the molecules are being cleaved by P1 nuclease (Figure 6).

Formation of *H-DNA is facilitated by negative supercoiling. When relaxed SV/CT$_{22}$ DNA was modified with OsO_4–pyridine and then digested with P1 as described before, no significant cleavage was observed even at the highest $ZnCl_2$ concentration tested (32 mM) (Figure 6), indicating that transition to *H-DNA of the d(GA.CT)$_{22}$ sequence at neutral pH requires negative superhelicity.

Discussion

In this paper we have presented data that indicate that the homopurine–homopyrimidine alternating d(GA.CT)$_{22}$ sequence adopts at neutral pH, in the presence of zinc ions, an altered DNA conformation (*H-form), which differs from either right-handed B-DNA or the protonated H-form DNA.

*H-DNA is characterized by peculiar DEPC and OsO_4 modification patterns. Right-handed B-DNA is insensitive to either DEPC or OsO_4. On the other hand, H-DNA shows extensive DEPC modification at the 5′ half of the homopurine strand and specific OsO_4 modification of thymines localized at the centre of the sequence (Hanvey *et al.*, 1988; Johnston, 1988; Vojtišková *et al.*, 1988; Voloshin *et al.*, 1988). Conversely, it is the central region of the homopurine strand that is being preferentially modified by DEPC in *H-DNA, while OsO_4 modification occurs principally at thymines on the 3′ half of the sequence. *H-DNA does not seem to require base protonation and is stabilized by negative supercoiling.

From the present data, it is difficult to ascertain unambiguously the actual structural conformation adopted by the d(GA.CT)$_{22}$ sequence in the presence of zinc. The patterns of DEPC and OsO_4 modification obtained at acidic pH in the absence of zinc are consistent with a triplex model for the protonated H-form of the sequence (Lyamichev *et al.*, 1985). Modification patterns obtained at neutral pH in the presence of zinc are somehow opposite to those corresponding to H-DNA, suggesting that *H-DNA might be formed by a triplex where the 5′ region of the homopurine strand folds back upon itself, leaving the 3′ region of the homopyrimidine strand unpaired (Figure 7). Given the extent of OsO_4 modification observed, only about one-quarter of the pyrimidine strand, from thymine 1 to thymine 11, will be involved in such a triplex. This model is consistent with the pattern of DEPC modification observed for the homopurine strand. Such a triplex will be stabilized by GGC and AAT base triads. GGC as well as AAU triplets are known to be stable at neutral pH (Marck *et al.*, 1978; Broitman *et al.*, 1987). Our results show that from the two possible isomeric homopurine–homopurine–homopyrimidine triplexes, only

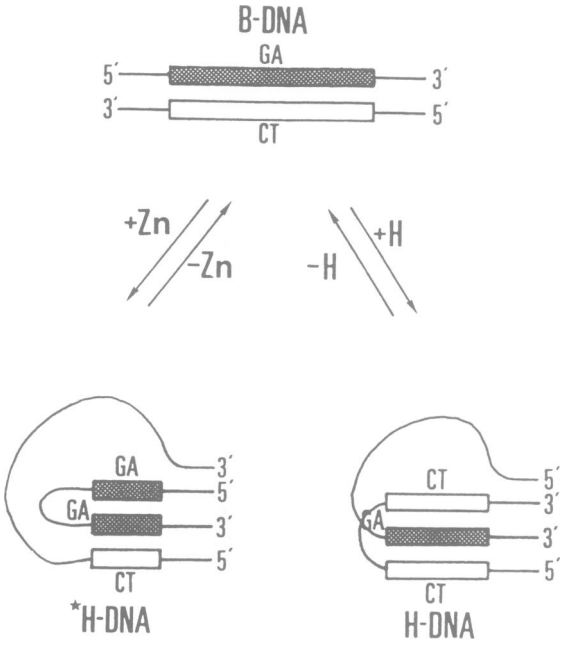

Fig. 7. Summary of the structural transitions described in this paper. At neutral pH in the absence of zinc, the $d(GA.CT)_{22}$ sequence exists predominantly as B-DNA. Increasing proton concentration results in a B-to-H transition. The patterns of chemical modification obtained at acidic pH in the absence of zinc are compatible with a triplex model for the protonated H-form. Increasing zinc concentration results in the formation of *H-DNA. The patterns of chemical modification obtained at neutral pH in the presence of 4 mM $ZnCl_2$ are consistent with a purine–purine–pyrimidine triplex, as schematically represented. At acidic pH in the presence of zinc, a mixture of both H- and *H-form is detected.

the one involving the 5′ part of the homopyrimidine strand is formed in the presence of zinc. Recently, a similar structural transition has been reported for $d(G.C)_n$ sequences, but induced by magnesium (Kohwi and Kohwi-Shigematsu, 1988). *H-DNA is also stabilized by other divalent cations but not magnesium (data not shown). A homopurine–homopyrimidine–homopurine triplex will also be compatible with our results. However, ATA and GCG base triads have not been observed to occur in solution (Broitman *et al.*, 1987). At present, it is unclear what is the role played by zinc ions in the stabilization of *H-DNA. They might actually interact specifically with the bases, in particular with the N-7 group of guanines (Daune, 1974; Tu and Heller, 1974). More work is required before establishing the type of ion–DNA interaction that results in the stabilization of *H-DNA.

Other structural models for *H-DNA are also possible. In fact, some of our observations, such as the dinucleotide repeat pattern of S1 nicking, the suppression of DEPC reactivity of guanines or the pattern of S1 sensitivity of the homopyrimidine strand, are not completely explained by a triplex model.

Figure 7 summarizes the different structural transitions described in this paper. At neutral pH in the absence of zinc ions, the $d(GA.CT)_{22}$ sequence exists principally as right-handed B-DNA as is indicated by its insensitivity to OsO_4 modification (Figure 5B). Acidification results in a B-to-H transition and, at pH 4.5, the $d(GA.CT)_{22}$ sequence exists mainly as H-DNA. On the other hand, increasing zinc concentration induces transition from B-DNA to *H-DNA. Formation of *H-DNA at neutral pH is induced by as little as 0.4 mM zinc (Figure 6). That *H-DNA is also present at acidic pH is shown by the OsO_4 modification pattern obtained at pH 4.5 in the presence of zinc (Figure 5B). However, the pattern of OsO_4 modification obtained at pH 4.5 in the presence of zinc is slightly different from that obtained at pH 7.0. In particular, thymines at the centre of the sequence appear to be more hyperreactive at acidic pH than at neutral pH, suggesting that at pH 4.5 in the presence of 4 mM $ZnCl_2$ both conformations are actually present, although the equilibrium is displaced to *H-DNA.

DEPC modification at pH 4.5 in the presence of zinc failed to detect formation of *H-DNA (Figure 5A). Moreover, the pattern of DEPC reactivity characteristic of H-DNA is also observed at neutral pH in the absence of zinc (Figure 5A). Similar results have been interpreted by others as indicative of the formation of a novel conformation (J-DNA) at neutral pH (Htun and Dahlberg, 1988). However, maintenance of DEPC reactivity at neutral pH is likely to reflect the fact that DEPC and OsO_4 modifications are carried out under slightly different experimental conditions which might influence the equilibrium between the various conformations. Under DEPC modification conditions, H-DNA appears to be more stable than under OsO_4 modification conditions. In fact, addition of DEPC results in a slight acidification of the medium that in our case (50 mM triethanolamine) dropped from pH 7.0 to pH 6.6.

Alternating homopurine–homopyrimidine $d(GA.CT)_n$ sequences are hypersensitive to cleavage by S1 nuclease. As we have shown here, the S1 hypersensitivity of the sequence does not depend significantly on pH. In general, S1 digestion experiments are carried out in the presence of zinc ions at concentrations that are sufficient to induce formation of *H-DNA within the homopurine–homopyrimidine sequence. S1 hypersensitivity of the $d(GA.CT)_{22}$ sequence at neutral pH is likely to reflect formation of *H-DNA, since this is the major altered conformation that the sequence adopts under these ionic conditions. Furthermore, S1 cleavage of the $d(GA.CT)_{22}$ sequence at neutral pH depends on supercoiling (Figure 3), as does formation of *H-DNA (Figure 6). In agreement with this interpretation, the S1 nicking pattern of the homopurine strand at pH 7.0 coincides quite well with the pattern of DEPC modification obtained at this pH in the presence of zinc, as would be expected if S1 was sensing principally *H-form DNA under these conditions. On the other hand, at acidic pH in the presence of zinc, a mixture of both H and *H-DNA is detected. If both conformations are sensitive to S1 nuclease, the S1 nicking pattern at pH 4.5 should be the sum of the nicking patterns due to each conformation. According to the DEPC modification data, the 5′ half of the homopurine strand should be nicked principally in H-DNA. Therefore, a displacement of the nicking pattern, as observed in Figure 4, towards the 5′ half of the homopurine strand would be expected if, under these conditions, S1 nuclease is actually sensing both conformations. In general, nicking of the homopurine strand at acidic pH occurs principally at the 5′ half of the sequence but displaced to the centre (Hanvey *et al.*, 1988).

The S1 nicking patterns obtained for the homopyrimidine strand are difficult to explain in terms of triplex models. Neither an H-triplex, which should be principally nicked at the centre, nor an *H-triplex, which should be principally nicked at the 3′ half, can account for the S1 hypersensitivity of the 5′ half of the homopyrimidine strand. In this case, S1 nicking occurs principally at the junction region. It might be that S1 is detecting some peculiar structural alteration ocurring on this region. On the other hand, if *H-DNA is actually a triplex, as the one schematically represented in Figure 7, the S1 insensitivity of the 3′ region of the homopyrimidine strand would suggest that the unpaired region of the pyrimidine strand is closely associated with the triplex being sterically inaccessible to S1 nuclease. Unlike S1 nuclease, OsO_4 would not experience such steric hindrance since it is a much smaller reagent.

The results shown here indicate that homopurine−homopyrimidine sequences, and in particular d(GA.CT)$_n$ sequences, are highly polymorphic from a structural point of view. Depending upon the precise ionic environment they can exist under different structural conformations. Addition of zinc ions induces transition to *H-DNA of a d(GA.CT)$_{22}$ sequence. On the other hand, magnesium ions mediate a transition from a CGC$^+$ triplex to the GCG triplex in d(G.C)$_n$ sequences (Kohwi and Kohwi-Shigematsu, 1988). Moreover, on the basis of circular dichroism data, Antao *et al.* (1988) have identified six different acid-induced conformations of linear poly d(GA.CT)$_n$ at pH values between 8 and 2.5. These different structural rearrangements include not only the poly d(C$^+$T.GA.CT)$_n$ triplex but also other protonated structures such as a self-complexed form of the polypurine strand or a loop-out structure of the polypyrimidine strand. In addition, Parniewski *et al.* (1989) have reported that, depending on pH and negative superhelicity, short d(GA.CT)$_n$ sequences can adopt at least two different non-B conformations. The high degree of structural versatility shown by homopurine−homopyrimidine sequences is likely to be responsible for the discrepancies between the different models proposed for their altered conformation(s).

Homopurine−homopyrimidine sequences are likely to play an important biological role in eukaryotes. They are frequently found at interesting locations in the genome: upstream of many eukaryotic genes (Larsen and Weintraub, 1982; Nickol and Felsenfeld, 1983; Schon *et al.*, 1983; Elgin, 1984; Evans *et al.*, 1984; McKeon *et al.*, 1984; Kilpatrick *et al.*, 1986), in the vicinity of replication origins (Soeda *et al.*, 1979; Riley *et al.*, 1986; Mirkin *et al.*, 1987) or in sites involved in genetic recombination (Davis *et al.*, 1980; Hentschel, 1982; Moos and Gallwitz, 1983; Htun *et al.*, 1984; Wohlrab *et al.*, 1987). Moreover, a nuclear protein with nucleolitic activity has been described to interact specifically with d(G.C)$_n$ sequences in eukaryotic chromatin (Ruiz-Carrillo and Renaud, 1987). Finally a d(GA.CT)$_n$ sequence cloned into SV40 was found to affect its *in vivo* replication rate (Rao *et al.*, 1988). The high degree of structural polymorphism shown by homopurine−homopyrimidine sequences, and in particular its ability to form triple-stranded structures, might be biologically relevant. Formation of a triplex within such sequences will actually open up the double helix, providing a single-stranded region which, depending on the precise triplex formed, might be used by factors that mediate such processes.

Materials and methods

DNAs

Recombinant SV/CT$_{22}$ and SV/O viruses have been described in detail elsewhere (Casasnovas *et al.*, 1987). Viruses were propagated in CV1 cells in Dulbecco's modified Eagle's medium (GIBCO-BRL) supplemented with 5% fetal calf serum (GIBCO-BRL). DNAs were prepared at 44 h post-infection according to Hirt (1967) and purified by centrifugation through ethidium bromide−CsCl gradients. Purified DNAs were stored in 10 mM Tris−HCl, pH 7.6, 1 mM EDTA at 4°C. Relaxed DNAs were prepared by treatment with DNA topoisomerase-I (BRL).

S1 nuclease digestion experiments

To determine the hypersensitivity to S1 nuclease cleavage of the d(GA.CT)$_{22}$ sequence, 100 ng of either SV/CT$_{22}$ or SV/O DNA were treated with S1 nuclease (SIGMA) in 10 μl of 50 mM NaCl, 1.5 mM $ZnCl_2$ at the following pH values: pH 4.5 and pH 5.0 (50 mM sodium acetate); pH 5.5 and pH 6.0 (50 mM bis-Tris) and pH 7.0 (50 mM triethanolamine). Digestions were performed at the following enzyme/DNA ratios: 0.1−1 U/μg at pH 4.5; 1−10 U/μg at pH 5.0 and pH 5.5; 10−50 U/μg at pH 6.0 and 10−100 U/μg at pH 7.0. All digestions were carried out at 15°C for 30 min. Digestions were stopped by the addition of 0.1 vol of 0.2 M EDTA, pH 7.0, and samples, after neutralization with 1 M Tris when necessary, were loaded onto a 1% agarose−TPE gel. Experiments were always performed in triplicate. For quantitation, ethidium bromide stained gels were photographed and the negatives scanned densitometrically in a Joyce−Loelb densitometer. The extent of S1 digestion was calculated as the ratio of nicked versus total DNA in each track corrected for the nicked DNA present in the untreated samples. Hypersensitivity was calculated as the ratio of the S1 sensitivities of SV/CT$_{22}$ versus SV/O.

Mapping of S1-hypersensitive sites

Gross-mapping of S1-hypersensitive sites in SV/CT$_{22}$ DNA was obtained by digestion of 0.4 μg DNA with 25 U S1/μg DNA at pH 4.5 and 100 U S1/μg DNA at pH 7.0 in the buffers and conditions described above. Digestions were terminated by phenol extraction and ethanol precipitation. S1-treated DNAs were then subjected to restriction cleavage with either *Eco*RI or *Taq*I restriction endonucleases (Boehringer) and analysed in 1% agarose−TBE gels.

For fine mapping of S1-hypersensitive sites, SV/CT$_{22}$ DNA was digested with S1 nuclease under conditions where no linear DNA was produced. Then, DNAs were cleaved at the unique *Ban*I site and either filled in with [α-^{32}P]dATP (Amersham) and the Klenow fragment of DNA polymerase I (Boehringer) or 5′ end-labelled with [γ-^{32}P]ATP (Amersham) and polynucleotide kinase (NEN). Subsequently DNAs were cleaved at the unique *Eco*RI site and the fragment of interest purified and analysed on a 6% polyacrylamide sequencing gel.

OsO_4 and DEPC modifications

For OsO_4 modification, intact supercoiled SV/CT$_{22}$ DNA (2−3 μg) was reacted with 1 mM OsO_4, 1% pyridine in a final volume of 50 μl for 30 min at 25°C in a buffer containing either 50 mM NaCl, 50 mM acetate, pH 4.5 or 50 mM NaCl, 50 mM triethanolamine, pH 7.0 and different amounts of $ZnCl_2$, from 0.1 mM to 32 mM, when indicated. Reactions were terminated by two successive ethanol precipitations. To determine the extent of OsO_4 modification, DNAs were linearized with *Eco*RI endonuclease and subsequently digested with P1 nuclease (Boehringer) in 50 mM NaCl, 10 mM $MgCl_2$, 10 mM Tris−HCl, pH 8.0. For fine mapping of the OsO_4 modification sites, DNAs were opened at the unique *Ban*I site and labelled, digested and purified as indicated above. Strand scission at modified pyrimidines was performed in 100 μl by treatment with 1 M piperidine at 90°C for 25 min. DNAs were then precipitated and freeze-dried twice.

For DEPC modification, intact supercoiled SV/CT$_{22}$ DNA (2−3 μg) was modified with 3 μl of DEPC (Sigma) for 15 min at 37°C in the same buffers as above. Reactions were terminated by ethanol precipitation as above. DNAs were then opened at the unique *Ban*I site and labelled, digested and processed as indicated above.

Modified DNAs were analysed on 6% polyacrylamide sequencing gels.

Acknowledgements

We are grateful to Drs J.A.Subirana and L.Cornudella for advice and support. This work was financed by grants from the CICYT (BIO88-0236), from the CEE (BAP-R5B), and ST2J-0372-C(A) and the FIS (88/1447). J.B. was a recipient of a postdoctoral fellowship from CSIC, R.B. and J.M.C. were recipients of doctoral fellowships from MEC.

References

Antao,V.P., Gray,D.M. and Ratliff,R.L. (1988) *Nucleic Acids Res.*, **16**, 719–738.
Broitman,S.L., Im,D.D. and Fresco,J.R. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5120–5124.
Cantor,C.R. and Efstratiadis,A. (1984) *Nucleic Acids Res.*, **12**, 8059–8072.
Casasnovas,J.M., Ellison,M.J., Rodriguez-Campos,A. and Azorín,F. (1987) *Eur. J. Biochem.*, **167**, 489–492.
Daune,M. (1974) In *Metal Ions in Biological Systems*. Marcel Dekker, New York, Vol. 3, pp. 1–43.
Davis,M.M., Kim,S.J. and Hood,L.E. (1980) *Science*, **209**, 1360–1365.
Elgin,S.C.R. (1984) *Nature*, **309**, 213.
Evans,T. and Efstratiadis,A. (1986) *J. Biol. Chem.*, **261**, 14771–14780.
Evans,T., Schon,E., Gora-Maslak,G., Patterson,J. and Efstratiadis,A. (1984) *Nucleic Acids Res.*, **12**, 8403–8058.
Galazka,G., Palecek,E., Wells,R.D. and Klysik,J. (1986) *J. Biol. Chem.*, **261**, 7093–7098.
Gross,D.S. and Garrard,W.T. (1986) *Mol. Cell. Biol.*, **6**, 3010–3013.
Hanvey,J.C., Klysik,J. and Wells,R.D. (1988) *J. Biol. Chem.*, **263**, 7386–7396.
Hentschel,C.C. (1982) *Nature*, **295**, 714–716.
Herr,W. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 8009–8013.
Hirt,B. (1967) *J. Mol. Biol.*, **26**, 365–369.
Htun,H., Lund,E. and Dahlberg,J.E. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7288–7292.
Htun,H. and Dahlberg,J.E. (1988) *Science*, **241**, 1791–1796.
Johnston,B.H. (1988) *Science*, **241**, 1800–1804.
Johnston,B.H. and Rich,A. (1985) *Cell*, **42**, 713–724.
Kilpatrick,M.W., Torri,A., Kang,D.S., Engler,J.A. and Wells,R.D. (1986) *J. Biol. Chem.*, **261**, 11350–11354.
Kohwi,Y. and Kohwi-Shigematsu,T. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 3781–3785.
Kohwi-Shigematsu,T. and Kohwi,Y. (1985) *Cell*, **43**, 199–206.
Larsen,A. and Weintraub,H. (1982) *Cell*, **29**, 609–622.
Lilley,D.M.J. and Palecek,E. (1984) *EMBO J.*, **3**, 1187–1192.
Lukasova,E., Jelen,F. and Palecek,E. (1982) *Gen. Physiol. Biophys.*, **1**, 53–70.
Lyamichev,V.J., Mirkin,S.M. and Frank-Kamenetskii,M.D. (1985) *J. Biomol. Struct. Dyn.*, **3**, 327–338.
Lyamichev,V.J., Mirkin,S.M. and Frank-Kamenestkii,M.D. (1986) *J. Biomol. Struct. Dyn.*, **3**, 667–669.
Lyamichev,V.J., Mirkin,S.M. and Frank-Kamenetskii,M.D. (1987) *J. Biomol. Struct. Dyn.*, **5**, 275–282.
Marck,C., Thiele,D., Schneider,C. and Guschlbauer,W. (1978) *Nucleic Acids Res.*, **5**, 1979–1996.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.
McKeon,C., Schmidt,A. and de Combrugghe,B. (1984) *J. Biol. Chem.*, **259**, 6636–6640.
Mirkin,S.M., Lyamichev,V.I., Drushlyak,K.N., Dobrynin,V.N., Filippov,S.A. and Frank-Kamenetskii,M.D. (1987) *Nature*, **330**, 495–496.
Moos,M. and Gallwitz,D. (1983) *EMBO J.*, **2**, 757–761.
Nejedly,K., Kwinkowski,M., Galazka,G., Klysik,J. and Palecek,E. (1985) *J. Biomol. Struct. Dyn.*, **3**, 467–478.
Nickol,J.M. and Felsenfeld,G. (1983) *Cell*, **35**, 467–477.
Parniewski,P., Galazka,G., Wilk,A. and Klysik,J. (1989) *Nucleic Acids Res.*, **17**, 617–627.
Pulleyblank,D.E., Haniford,D.B. and Morgan,A.R. (1985) *Cell*, **42**, 271–280.
Rao,B.S,. Manor,H. and Martin,R.G. (1988) *Nucleic Acids Res.*, **16**, 8077–8094.
Riley,D.E., Reeves,R. and Gartler,S.M. (1986) *Nucleic Acids Res.*, **14**, 9407–9423.
Ruiz-Carrillo,A. and Renaud,J. (1987) *EMBO J.*, **6**, 401–407.
Soeda,E., Arrand,J.R., Smolar,N. and Griffin,B.E. (1979) *Cell*, **17**, 357–370.
Schon,E., Evans,T,. Welsh,J. and Efstratiadis,A. (1983) *Cell*, **35**, 838–848.
Tautz,D. and Renz,M. (1984) *J. Mol. Biol.*, **172**, 229–235.
Tu,T. and Heller,M.J. (1974) In *Metal Ions in Biological Systems*. Marcel Dekker, New York, Vol. 1, pp. 2–49.
Vogt,V.M .(1980) *Methods Enzymol.*, **65**, 248–255.
Vojtišková,M. and Palecek,E. (1987) *J. Biomol. Struct. Dyn.*, **5**, 283–296.
Vojtišková,M, Mirkin,S., Lyamichev,V., Voloshin,O., Frank-Kamenetskii,M. and Palecek,E. (1988) *FEBS Lett.*, **234**, 295–299.
Voloshin,O.N., Mirkin,S.M., Lyamichev,V.I., Belotserkovskii,B.P. and Frank-Kamenetskii,M.D. (1988) *Nature*, **333**, 475–476.
Weintraub,H. (1983) *Cell*, **32**, 1191–1203.
Wohlrab,F., McLean,M.J. and Wells,R.D. (1987) *J. Biol. Chem.*, **262**, 6407–6416.

Received on February 22, 1989; revised on April 20, 1989